Inferencia Estadística II

Junio de 2004

Problema 1 (3 puntos)

1. El test de Neyman-Pearson (NP) está basado en el ratio de verosimilitudes :

$$\frac{f_1(X)}{f_0(X)} = \frac{5e^{5(X-1)}}{\frac{1}{2}}$$

Claramente este ratio crece con X, por lo tanto el test de NP se puede expresar como

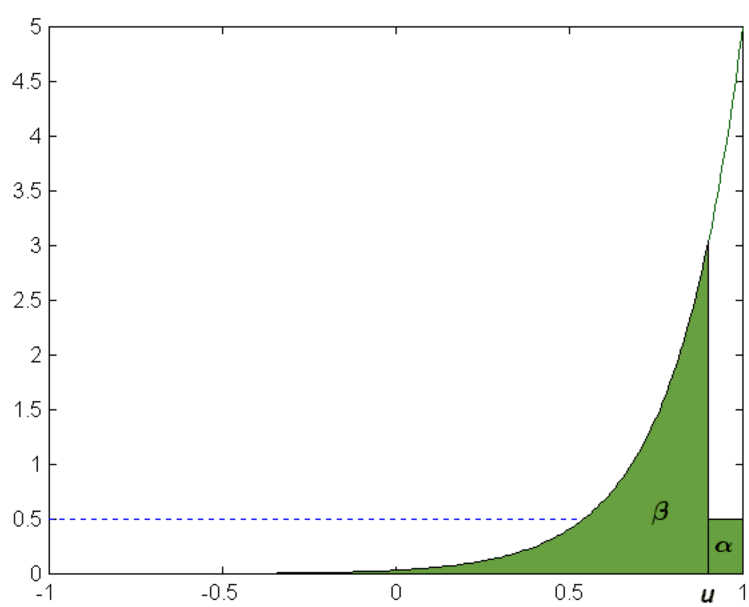
$$\text{Rechazar } H_0 \text{ si } X > u$$

El nivel del test $\alpha = 5\%$, entonces el umbral u debe verificar

$$\begin{aligned} 5\% &= \mathbf{P}\left(X > u | H_0\right) \\ &= \frac{1}{2}\int_u^1 dx = \frac{1}{2}\left(1-u\right) \end{aligned}$$

De la anterior ecuación se deduce que $u_\alpha = 0.9$.

2. En el gráfico a continuación aparece los dos riesgos α y β asociados al test de NP :



3. La potencia del test se define como la probabilidad de rechazar H_0 cuando es falsa, o sea

$$\begin{aligned} \mathbf{P}\left(X > u_\alpha | H_1\right) &= \int_{0.9}^1 5e^{5(x-1)}dx \\ &= \left[e^{5(x-1)}\right]_{0.9}^1 = 1 - e^{-0.5} \simeq 40\% \end{aligned}$$

Por lo tanto, sólo en el 40% de los casos el padre será capaz de detectar si la escopeta está trucada basándose en un único disparo.

Problema II (3 puntos). El test adecuado para contrastar la independencia entre ambas variables cualitativas (color de los ojos y color del pelo) está basado en el estadístico de Pearson :

$$\begin{aligned} K_n^* &= \sum_{i=1}^{3}\sum_{j=1}^{4} \frac{\left(N_{ij} - \frac{N_{i.}N_{.j}}{n}\right)^2}{\frac{N_{i.}N_{.j}}{n}} \\ &= \left[\frac{\left(25 - \frac{44\cdot 45}{124}\right)^2}{\frac{44\cdot 45}{124}} + \ldots + \frac{\left(8 - \frac{21\cdot 33}{124}\right)^2}{\frac{21\cdot 33}{124}}\right] \\ &\simeq 15.1 \end{aligned}$$

donde $n = 124$.
Sabemos que la distribución asintótica de este estadístico es una χ^2 con $(3-1)(4-1) = 6$ grados de libertad. El cuantil 5% de esta distribución siendo igual a $\chi^2_{6,5\%} = 12.59$, rechazamos la hipótesis de independencia.

Problema III (4 puntos)..

1. La función de perdida siendo cuadrática, la mejor predicción de θ será aquí :

$$d_Q = \mathbf{E}_Q(\theta) = 10$$

   puesto que la distribución *a priori* Q es exponencial con media $\frac{1}{\lambda} = 10$. El riesgo bayesiano de esta predicción es $\overline{w}_Q(d_Q) = \mathbf{var}_Q(\theta) = \frac{1}{\lambda^2} = 100$.

2. .

   (a) La distribución de θ condicionada por X es una Gamma $G(X+1, \lambda+1)$ con $\lambda = \frac{1}{10}$.

   (b) Por lo tanto, la mejor predicción de θ conociendo X será :

$$d_Q(X) = \mathbf{E}_Q(\theta|X) = \frac{X+1}{\lambda+1} = \frac{X+1}{1.1}$$

   (c) El riesgo de esta predicción es

$$\begin{aligned} W_Q(d_Q(X)) &= \mathbf{E}\left[\mathbf{var}_Q(\theta|X)\right] \\ &= \mathbf{E}\left[\frac{X+1}{(\lambda+1)^2}\right] \\ &= \frac{\mathbf{E}[X] + \mathbf{1}}{(\lambda+1)^2} = \frac{\lambda^{-1}+1}{(\lambda+1)^2} = \frac{1}{\lambda(\lambda+1)} \end{aligned}$$

   porque $\mathbf{E}[X] = \mathbf{E}[\mathbf{E}(X|\theta)] = \mathbf{E}[\theta] = \lambda^{-1}$.
   Puesto que $\lambda(\lambda+1) > \lambda^2$ concluimos que $W_Q(d_Q(X)) < \overline{w}_Q(d_Q)$, o sea que $d_Q(X)$ es más preciso que d_Q.